The EMBO Journal vol.9 no.5 pp.1603-1614, 1990

Two distinct estrogen-regulated promoters generate transcripts encoding the two functionally different human progesterone receptor forms A and B

P.Kastner, A.Krust, B.Turcotte, U.Stropp, L.Tora, H.Gronemeyer and P.Chambon

Laboratoire de Génétique Moléculaire des Eucaryotes, CNRS, Unité 184 de Biologie Moléculaire et de Génie Génétique, INSERM, Institut de Chimie Biologique, Faculté de Médicine, 11 rue Humann, 67085 Strasbourg Cédex, France

Communicated by P.Chambon

The human progesterone receptor (hPR) cDNA, synthesized from T47D breast cancer cells, and the hPR gene 5'-flanking region were cloned and sequenced. Comparison of the cDNA-deduced amino acid sequence with other PR homologues demonstrated the modular structure characteristic of nuclear receptors. As in the case of the chicken homologue, there are two hPR forms, A and B, which originate from translational initiation at AUG2 (codon 165) and AUG1, respectively. Northern blot analysis of T47D mRNA using various cDNA derived probes identified two classes of hPR mRNAs, one of which could code for hPR form B, while the other one lacked the 5' region upstream of AUG1. S1 nuclease mapping and primer extension analyses confirmed that the second class of hPR transcripts are initiated between +737 and +842 and thus encode hPR form A, but not form B. By using the hPR gene 5'-flanking sequences as promoter region in chimeric genes, we show that a functional promoter (located between −711 and +31) directs initiation of hPR mRNAs from the authentic start sites located at +1 and +15. Most importantly, initiation of transcription from chimeric genes demonstrated the existence of a second promoter located between +464 and +1105. Transient co-transfection experiments with vectors expressing the human estrogen receptor showed that both promoters were estrogen inducible, although no classical estrogen responsive element was detected in the corresponding sequences. When transiently expressed, the two hPR forms similarly activated transcription from reporter genes containing a single palindromic progestin responsive element (PRE), while form B was more efficient at activating the PRE of the mouse mammary tumor virus long terminal repeat. Transcription from the ovalbumin promoter, however, was induced by hPR form A, but not by form B.

Key words: alternative promoters/estrogen responsive element/progesterone receptor/progesterone responsive element/T47D cells

Introduction

The human progesterone receptor (hPR) belongs to the superfamily of nuclear receptors (including those for steroid and thyroid hormones, and retinoids) whose members co-ordinate morphogenesis and homeostasis in response to the binding of their ligands. These receptors act as transcriptional factors that regulate gene expression positively or negatively by interacting with cognate DNA sequences [ligand responsive elements (REs), enhancers in the case of positive regulation]. Extensive structure−function analyses have localized and characterized multiple domains which are contained in discrete, differentially conserved segments (designated A to F) of receptor primary structure (for reviews, see Evans, 1988; Green and Chambon, 1988; Gronemeyer *et al.*, 1988; Ham and Parker, 1989; Beato, 1989). The DNA binding domain (DBD, contained within region C) is thought to fold into so-called 'zinc-fingers' and only a very limited number of amino acids within the knuckle of each finger is decisive for specific recognition of the cognate RE (Freedman *et al.*, 1988; Danielsen *et al.*, 1989; Mader *et al.*, 1989; Umesono and Evans, 1989). The ligand-free steroid hormone binding domain (HBD, contained within region E) prevents the DBD from binding the RE and

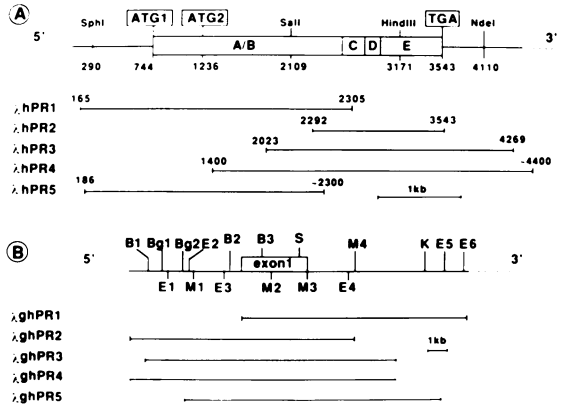
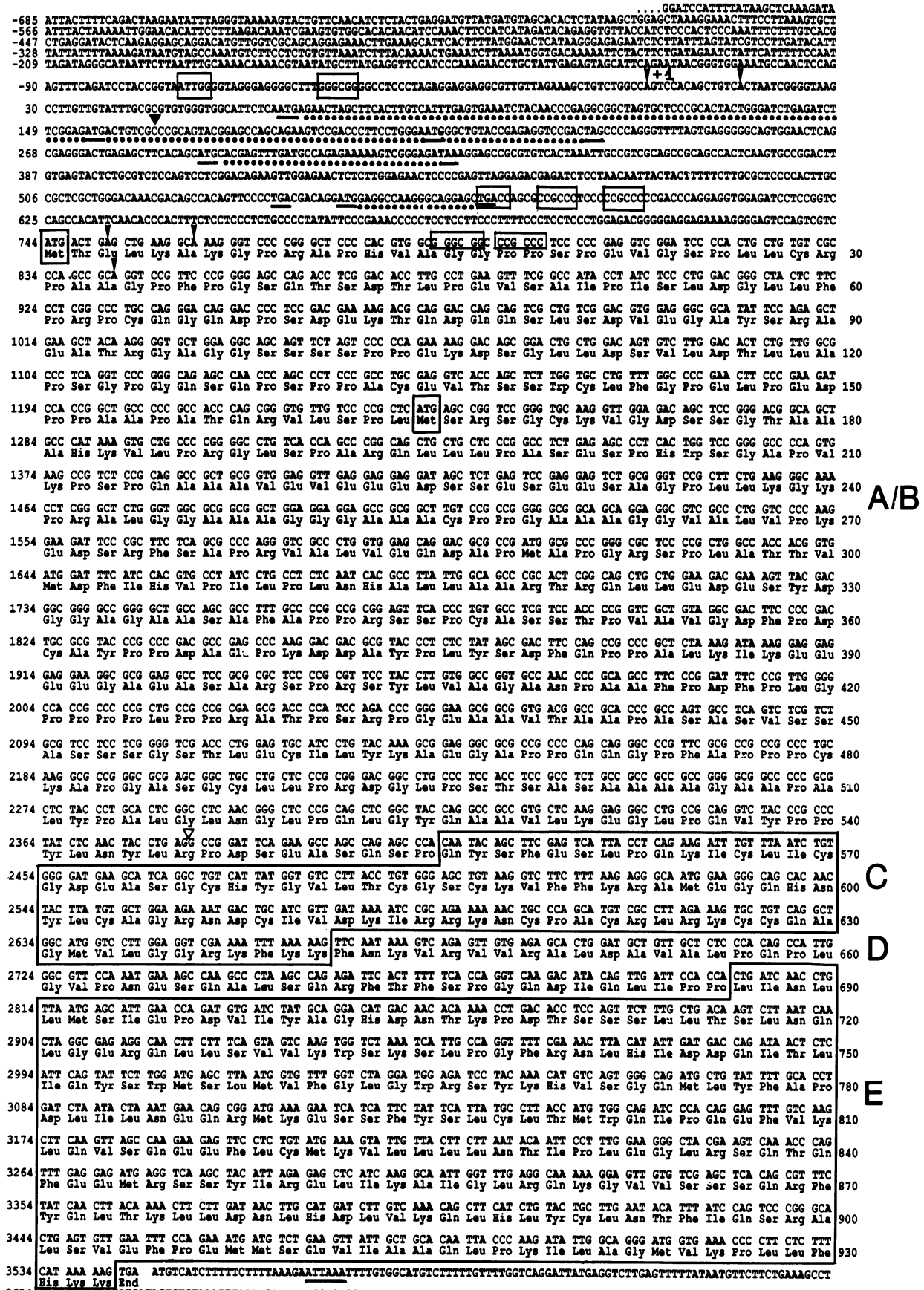


Fig. 1. Schematic representation of hPR cDNA and genomic clones. (**A**) λhPR1 to 4 were isolated from a T47D λgt11 cDNA library by screening with the chicken λcPR5 cDNA (Jeltsch *et al.*, 1986), while λhPR5 was obtained by rescreening the library with a fragment of λhPR4 spanning position 1400−2109 relative to the hPR cap site (see Figure 2). The position of the 5' and 3' extremity of each cDNA insert is given, the exact 3' ends of λhPR4 and λhPR5, however, have not been determined. At the top a scheme of the coding region is depicted, with boxes C and E indicating DNA and hormone binding domains. The N-terminal region A/B contains the two ATGs (ATG1 at +744; ATG2 at +1236) which are the translational start sites of hPR forms B and A, respectively (see text). The hPR stop codon (TGA) and, for orientation, some restriction sites are depicted with their positions according to the cDNA sequence in Figure 2.
(**B**) Representation of hPR genomic clones (λghPR1 to 5) isolated from a human genomic λEMBL3 library by screening with the B3−M2 cDNA probe. A restriction map is given at the top (B1−B3, *Bam*HI; Bg1 and Bg2, *Bgl*II; M1−M4, *Mst*II; E1−E6, *Eco*RI; S, *Sal*I; K, *Kpn*I) and the position of the first hPR gene exon is indicated as a box. The 5' boundary of exon 1 was mapped as described in the text, the 3' end was determined by sequencing of the splice donor site present in the M3−E4 fragment which was found to be conserved in the human and chicken homologues [Jeltsch *et al.* (1989); the intron-exon boundary (open triangle in Figure 2) was determined as 5'-AACTACCTGAG/ GTGAGGGCCCGGGACGG-3', with CCTGAGG as the *Mst*II site in cDNA (compare Figure 2) and gene (B.Turcotte, unpublished results)].


A/B
C
D
E

a constitutive trancription activating function is present in the non-conserved region A/B. Ligand binding induces receptor dimerization prior to DNA binding and creates an 'active surface', within the HBD, thus generating a second transcriptional activating function (Webster *et al.*, 1988, 1989; Hollenberg and Evans, 1988; Kumar and Chambon, 1988; Godowski *et al.*, 1988; Tora *et al.*, 1989).

While the rabbit homologue (rbPR) has been proposed to exist as only one form (Loosfelt *et al.*, 1984; Logeat *et al.*, 1985), chick oviduct and human T47D breast cancer cells contain two forms, A and B, of the chicken (cPR) and human (hPR) progesterone receptor. The results obtained by transient expression of recombinant cPR, hPR and mutants derived therefrom, have suggested that form A may lack the N-terminal 127 (cPR) or 164 (hPR) amino acids present in form B and originate by internal initiation from an in-frame AUG present in the mRNA encoding form B (Conneely *et al.*, 1987b, 1989; Gronemeyer *et al.*, 1987; Krett *et al.*, 1988; Kastner *et al.*, 1990). However, we have recently characterized the cPR gene and mRNAs and identified a 5′-truncated transcript which selectively encodes form A, thus indicating that the two forms may originate from different mRNAs (Jeltsch *et al.*, 1990).

We describe here the cloning and sequencing of the hPR cDNA and gene 5′-flanking region and demonstrate that forms A and B are encoded by separate mRNAs which are transcribed from two distinct promoters, both of which are under estrogen control. We also show that transcription from promoters of progesterone responsive genes is activated differentially by hPR forms A and B, and discuss the possibility that their relative abundance could play a role in the control of gene expression by progestins.

Results

The cDNA sequence encoding hPR forms A and B

Twenty positive hPR clones were isolated from a randomly primed T47D λgt11 cDNA library, using the previously described cDNA insert of λcPR5 (containing parts of the DBD and HBD of cPR, Jeltsch *et al.*, 1986) as a probe. Restriction mapping identified four overlapping clones (λhPR1 to 4, Figure 1A) whose inserts, covering ~4.2 kb, were sequenced (Figure 2, the most 5′ end of λhPR1 is indicated by a filled triangle at position +165; note that λhPR4 was only partially sequenced). Sequences of an independently derived clone (λhPR5, Figure 1A), identified by using the cDNA insert of λhPR4 as a probe, were also determined. Sequence analysis revealed a major open reading frame of 2799 nucleotides encoding a 933 amino acid protein (M_r 99054) with an upstream in-frame terminator codon (at +540 in Figure 2) and several upstream open reading frames (UORFs) in the 5′-untranslated region (see Discussion).

Alignment of the hPR amino acid sequence with those of the chicken (Gronemeyer *et al.*, 1987) and rabbit (Loosfelt *et al.*, 1986) homologues revealed (nomenclature as defined by Krust *et al.*, 1986) a 100% conserved region C and a highly conserved region E (see Figure 2) which have been shown to contain the DBD and HBD of cPR, respectively (Eul *et al.*, 1989; our unpublished experiments have confirmed these assignments for hPR). Regions A/B (Figure 3) and D are less conserved between the three species; 30% sequence identity was found for the chicken and human regions A/B. There are, however, several blocks of conserved sequences which may indicate a common function of these A/B regions (see below). Note also that in contrast to the cPR case (amino acids 49–77) there is no polyglutamic acid tract in hPR and rbPR and that the difference in size between hPR, rbPR and cPR is due to the different lengths of regions A/B (see Figures 2 and 3, and Gronemeyer *et al.*, 1987). The hPR cDNA sequence has also been reported by Misrahi *et al.* (1987); minor sequence differences with the one given here are listed in the legend to Figure 2.

We (Kastner *et al.*, 1990) and others (Krett *et al.*, 1988) have shown that hPR form A can be expressed from a vector (e.g. hPR2, see Figure 8A) generating transcripts from which translation is initiated at ATG2 (codon 165, boxed in Figure 2), while vectors (e.g. hPR0, hPR1, see Figure 8A) that generated transcripts containing ATG1 (codon 1, boxed in Figure 2) produced exclusively hPR form B when transfected into HeLa or Cos-1 cells (Kastner *et al.*, 1990; see also Discussion).

A promoter region is located at the 5′ end of the hPR form B transcription unit

To isolate clones containing the 5′ end of the hPR form B transcription unit, a λEMBL3-based human leucocyte genomic library (gift of Transgène SA, Strasbourg) was screened with the *Bam*HI–*Mst*2 fragment of λhPR1 (Figure 1A, +814 to +1105 in Figure 2). Five clones (λghPR1 to 5) were isolated and positioned relative to each other by restriction mapping and Southern blot analysis (Figure 1B and data not shown). The B2–B3 fragment of λghPR2 was sequenced (see Figure 2, −711 to +814) and found to be co-linear with the cDNA sequence described above.

S1 nuclease mapping and nuclear run-on assays were used to localize the hPR cap site (+1 site). An 'S1 probe'

Fig. 2. Nucleotide sequence of the hPR cDNA and promoter region, and deduced amino acid sequence of the hPR protein. The 5′ part of the cDNA sequence was determined from λhPR1; its 5′ end is indicated by a solid triangle. Sequences further upstream were obtained from the B2–B3 fragment of the genomic clone λghPR2. Numbers on the left refer to positions relative to +1 which corresponds to the most 5′ start site of T47D hPR mRNAs originating from promoter B (see text); numbers on the right refer to amino acid positions. Regions A/B to E (see text) are indicated. Region C is 100% conserved in hPR, rbPR and cPR, while region E of hPR has 99 and 87% sequence identity with the rabbit and chicken homologues, respectively. Bold boxes indicate the translation initiation codons of hPR form B (+744) and form A (+1236). An in-frame stop codon (at +540) upstream of the form B ORF is underlined. Major transcription start sites (at +1, +15 and at +751, +761, +842) are indicated by arrows. An open triangle refers to the position of the first intron. Five ATG codons in the 5′-untranslated region and their in-frame termination codons are underlined, the corresponding UORFs are indicated by dotted lines. In the 3′-untranslated region a potential polyadenylation signal site is underlined. Putative Sp1-binding sites are boxed, as well as a 'CCAAT' motif and a putative half-palindromic ERE. The following differences have been observed with the hPR cDNA sequence reported by Misrahi *et al.* (1987). (i) +1402 A was G in their study, giving GGC (Gly) instead of AGC (Ser). (ii) +1511-TTGTCCG (Cys–Pro) was reported as TGTCCCG (Val–Pro). (iii) +2723 T was G, giving GTG (Val) instead of TTG (Leu). (iv) +3057 T was reported as C, giving CAC instead of CAT (both His). (v) +3405 G was A, giving CAA instead of CAG (both Gln). In our case, sequences at these positions were determined from at least two distinct cDNA clones, except for positions +1402, which was derived from λhPR1.

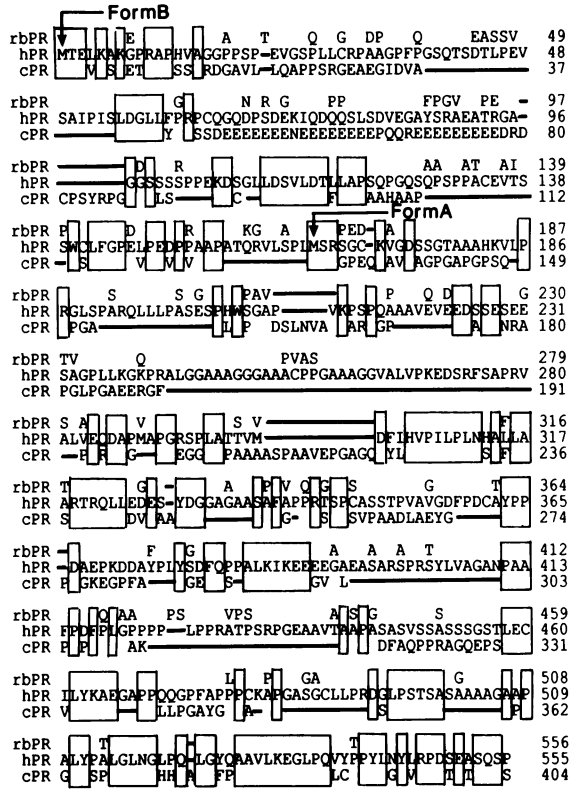


Fig. 3. Alignment of the A/B region of rbPR, hPR and cPR. Residues which are conserved in all three species are boxed. The human sequence is given in its entirety and only the divergent residues of cPR (Conneely *et al.*, 1987a; Gronemeyer *et al.*, 1987) and rbPR (Loosfelt *et al.*, 1986) are shown. Gaps introduced for optimal alignment are indicated by lines. The translational start sites of the hPR and cPR forms A and B are indicated by arrows.

(Figure 4A) was generated by primer extension and subsequent *Bam*HI restriction which, upon hybridization with T47D cell poly(A)$^+$ RNA and S1 nuclease digestion, yielded two protected fragments with their 5′ ends located 744 (+1 in Figure 2) and 730 (+15 in Figure 2) nucleotides upstream from the start of the hPR ORF (Figure 4B, compare lane 3 with the parallel sequencing gel, lanes 4–7). RNase protection assays identified the same sites (not shown). No protected fragments were seen with chick oviduct RNA (lane 1) and total T47D cell RNA gave only weak signals at +1 and +15 (lane 2). Moreover, while strand-specific transcripts were readily detected in nuclear run-on assays performed with single stranded templates derived from λghPR1 (Figure 1B; see below and Figure 7), no significant hybridization signal was obtained when using a single-stranded template derived from λghPR2 spanning position −711 to +31 (data not shown). Thus, we concluded that initiation of hPR gene transcription occurred at positions +1 and +15.

To demonstrate that sequences located upstream of +31 contain a functional promoter, we constructed chimeric recombinants bearing the hPR sequences from −711 to +31 in front of the promoterless bacterial chloramphenicol acetyl transferase (CAT) and rabbit β-globin genes, designated PR-[−711,+31]-CAT and PR-[−711,+31]-GLOB, respectively (illustrated in Figure 4A). Transient transfection of PR-[−711,+31]-CAT into HeLa cells resulted in CAT activity (Figure 4C, lanes 4–7), whereas no activity was observed with the promoterless parental pBLCAT3 (Figure 4C, lane 1). When analysed by S1 nuclease mapping, initiation of transcription occurred predominantly at +1 within the hPR sequences of hPR-[−711,+31]-GLOB (Figure 4D, lanes 4 and 5, upper open triangle; see the A track in lane 1 for localization in Figure 2). On overexposed autoradiographs two other sites of initiation at +15 (lower open triangle) and at +21 (solid triangle) were apparent (see also below). Note that this latter site was not observed in S1 nuclease mapping experiments using T47D mRNA (Figure 4B, lane 3). From the above results we concluded that the 5′ end of the transcription unit which gives rise to hPR form B is located 744 bp upstream of ATG1 and designated the corresponding promoter region as hPR promoter B.

Characterization of transcripts which encode only hPR form A and evidence that they are generated from a separate promoter A

To investigate whether distinct transcripts might encode the two forms of hPR, we performed S1 nuclease mapping and primer extension analyses to identify hPR mRNAs which may possibly be initiated between ATG1 and ATG2 (illustrated in Figure 5A). Using the 'S1 probe' spanning the *Bgl*II−*Nar*I fragment in Figure 5A, several protected fragments were observed after digestion by S1 nuclease of hybrids formed with T47D mRNA (Figure 5B, lane 5). To exclude that these S1 nuclease cleavage sites could be due to secondary structure features of hPR mRNA initiated at position +1, *in vitro* synthesized hPR RNA initiating at +1 was analyzed in parallel (Figure 5B, lane 6; see also Materials and methods). Clearly, four of the observed bands (asterisks in lane 5) appear to correspond to such 'false positive' S1 nuclease cleavage sites (note the potential of hPR RNA to form a stable hairpin between +687 and +733, see Discussion). While one minor 'non-artefactual' fragment may originate from transcription initiation immediately upstream of ATG1 at position +737 (filled triangle in lane 5; note the position of ATG1 from the parallel sequence ladder in lanes 1–4), three major protected fragments (open triangles in lane 5, arrows in Figure 2) indicated the existence of hPR transcripts with 5′-ends located downstream of ATG1 (+751, +761 and +842). No such fragments were obtained when using HeLa cell RNA (lane 7). Moreover, the same three sites were mapped by primer extension analysis (PRIMER in Figure 5A, data not shown), thus excluding the possibility that they correspond to splice sites (note also the absence of consensus splice acceptor sites at these positions in Figure 2). In conclusion, T47D cells apparently express a family of hPR mRNAs which possess 5′ ends starting downstream of ATG1 and, consequently, are able to generate hPR form A, but not form B.

To test whether sequences surrounding the 'downstream' start sites indeed exhibited promoter activity we constructed the chimeric gene PR-[464,1105]-CAT containing the hPR *Bgl*II−*Mst*II cDNA fragment in front of the promoterless CAT gene (Figure 5A). Significant CAT activity was observed when this chimeric gene was transfected into HeLa cells (Figure 5C, lanes 3–6), while no transcription was seen with the promoterless pBLCAT3 (lanes 7 and 8). From the above results we concluded that a second promoter generates the 'hPR form A' transcripts and designated the corresponding region as hPR promoter A.

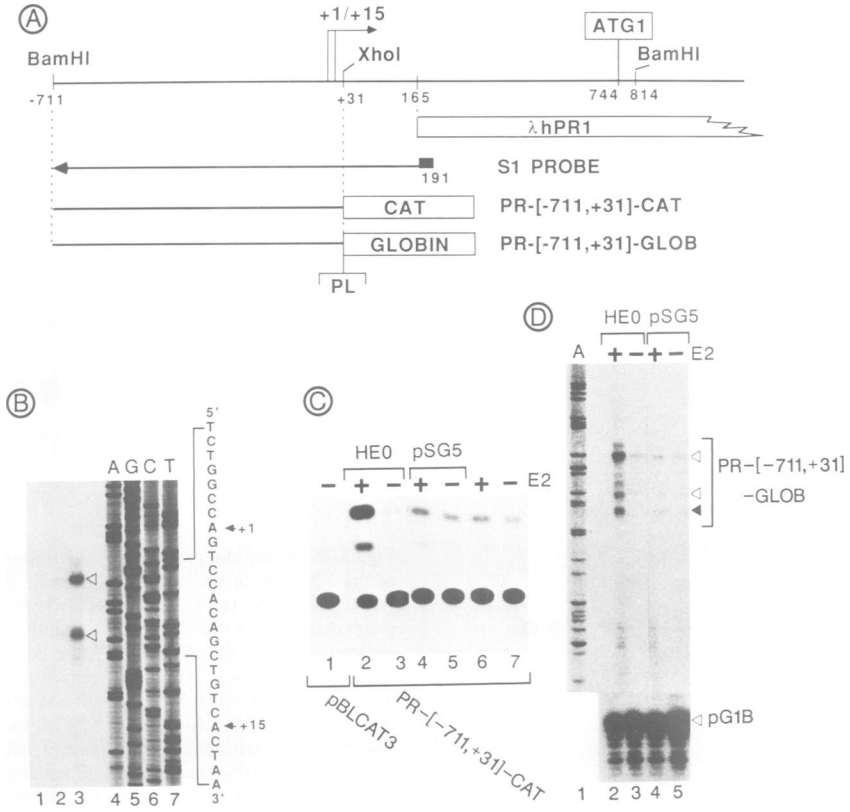


Fig. 4. Characterization of the distal hPR promoter B. (**A**) Schematic illustration of the B2−B3 (Figure 1B) region of the hPR gene indicating the two distal cap sites (+1 and +15), ATG1 and a *Xho*I site at +31 which was introduced by site-directed mutagenesis. The most 5′-extending cDNA clone (λhPR1) is depicted, as well as the oligonucleotide (complementary to nucleotides +172 to +191) used for priming the synthesis of the 'S1 probe'. The S1 probe was restricted by *Bam*HI and purified on a strand-separating gel (Maniatis *et al.*, 1982). The reporter genes hPR-[−711,+31]-CAT and hPR-[−711,+31]-GLOB, containing the same DNA fragment in front of promoterless CAT or globin genes, respectively, are schematically illustrated below. PL is a polylinker introduced during construction of the globin reporter gene (see Materials and methods). (**B**) Determination of upstream hPR mRNA cap sites by S1 nuclease mapping: 10 μg of chicken oviduct total RNA (lane 1), 10 μg of T47D total (lane 2), or 10 μg of T47D poly(A)$^+$ RNA (lane 3) were hybridized to the labelled S1 probe, treated with S1 nuclease and the resistant hybrids (open triangles) were separated on sequencing gels as described in Materials and methods. The corresponding DNA sequence run in parallel on the same gel is shown for orientation (lanes 4−7). The sequence shown on the right corresponds to the coding strand and was determined with the same primer as that used for probe synthesis. The two cap sites (+1, +15) are shown. (**C**) Transcriptional activation from promoter B is stimulated by the E2/hER complex: 5 μg of PR-[−711,+31]-CAT (lanes 2−7) or 5 μg of pBLCAT3 (lane 1) were transfected into HeLa together with 5 μg of HEO (lanes 2 and 3) or 5 μg of pSG5 (lanes 4 and 5). In all cases, 1 μg of β-galactosidase reference plasmid pCH110 (Pharmacia) was included in the co-transfection mix, and amounts of extracts with identical β-galactosidase activity were used to perform CAT assays. Where indicated, 17β-estradiol (E2) was added to the transfected cells, to give a final concentration of 10^{-8} M. (**D**) Initiation of transcription occurs at the authentic cap sites in chimeric genes bearing hPR promoter B: 5 μg of PR-[−711,+31]-GLOB and 500 ng of the internal control recombinant pG1B (Sassone-Corsi *et al.*, 1985) were transfected into HeLa cells together with 5 μg of HEO (lanes 2 and 3) or 5 μg of pSG5 (lanes 4 and 5). Cytoplasmic RNA was extracted and subjected to S1 nuclease mapping. Transcriptional start sites in pG1B (open triangle, bottom panel) and in the PR promoter chimeric gene (triangles, upper panel) are indicated. The solid triangle points to a protected fragment which was not observed with T47D poly(A)$^+$ RNA. The A-specific sequence ladder (lane 1) corresponds to the sense strand.

To correlate the above S1 nuclease mapping data with individual mRNA species and to determine their relative steady-state levels, Northern blotting experiments were done with T47D cell poly(A)$^+$ RNA using various oligonucleotide probes encompassing the 5′ region of the hPR transcription unit (probes A−D illustrated in Figure 6). Using either the entire hPR cDNA or a probe located downstream of ATG2 (e.g. probe D), six RNA species with estimated sizes of 11.4, 6.1, 5.2, 4.5, 3.7 and 2.9 kb were detected (Figure 6, lane 4 and data not shown; see also Read *et al.*, 1988; Wei *et al.*, 1988), indicating that all hPR mRNA species initiate upstream of ATG2. Since probe C produced a similar pattern to probe D, the corresponding promoters must be located upstream of +842 (lane 3). However, when using probes A or B, only a subset of the bands described above was detected (lanes 1 and 2). These corresponded to the 11.4, 6.1, 4.5 and 3.7 kb RNA species, while the 5.2 and 2.9 kb transcripts could not be seen even on overexposed blots (not shown). While the smallest mRNA species initiating upstream of ATG1 and thus having the potential to encode hPR form B (3.7 kb) apparently originates from polyadenylation at the signal site ATTAAA (underlined in Figure 2), the 4.5, 6.1 and 11.4 kb species may be the result of alternative polyadenylation (see Jeltsch *et al.*, 1990). Interestingly, the 'downstream' initiated mRNA species detected by probes C and D (2.9 and 5.2 kb, Figure 6) are 800−900 bp smaller than the 3.7 and 6.1 kb species. Since this difference is that expected between classes of transcripts initiating at either +1 and +15, or at the downstream cluster of start sites (+737 to +842), our data suggest that the 2.9 and 3.7 kb RNA species result from polyadenylation at the above described signal site, while a polyadenylation signal

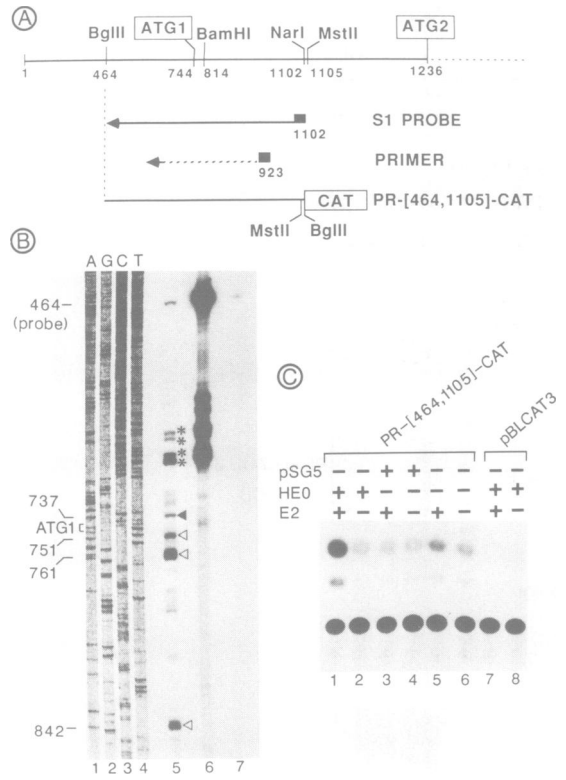


Fig. 5. Characterization of the proximal hPR promoter A. (**A**) A schematic representation of cDNA sequences upstream of ATG2 showing the positions of ATG1, and several restriction sites for orientation, is given at the top. The 'S1 probe', synthesized by extending a primer (+1086 to +1102; note that only probes upstream of +1102 were used for S1 nuclease mapping, since on Northern blots all hPR mRNA species were detected with *Bam*HI–*Nar*I probe, see Figure 6, lanes 3 and 4, and data not shown) and restricting with *Bgl*II before isolation from a strand-separating gel; the 'primer' (+888 to +932) used for primer extension with T47D poly(A)$^+$ RNA (see text), and the reporter recombinant PR-[+464,+1105]-CAT are depicted below. *Mst*II and *Bgl*II sites used in the construction of the reporter recombinant are indicated. (**B**) S1 nuclease mapping identifies a cluster of potential hPR mRNA cap sites located around ATG1: 10 μg of T47D poly(A)$^+$ RNA (lane 5), 50 ng of *in vitro* synthesized (Materials and methods) hPRO mRNA (Kastner *et al.*, 1990; Figure 8A), mixed with 10 μg of HeLa cell total RNA (lane 6), or 10 μg of HeLa cell total RNA alone (lane 7) were hybridized with the S1 probe, treated with S1 nuclease and resistant hybrids were separated on sequencing gels. Asterisks indicate hybrids which may be resistant to S1 nuclease due to RNA secondary structures (see text); triangles point to 'true' S1 nuclease signals, with open triangles indicating the putative cap sites of those hPR mRNAs that have no potential to encode hPR form B, but can generate form A. A parallel sequence ladder (lanes 1–4) is shown, with the positions of ATG1, cap sites (+737, +751, +761, +842) and probe 3′ end (+464), indicated. (**C**) Initiation of estrogen-inducible transcription from a chimeric recombinant containing hPR promoter A: 5 μg of PR-[+464, +1105]-CAT or of the promoterless vector pBLCAT3 were transfected into HeLa cells together with 1 μg HE0 or pSG5. Estradiol (E2) was added to cells at a final concentration of 10^{-8} M. Standardization (from co-transfected pCH110) and CAT assays were performed as described in Materials and methods.

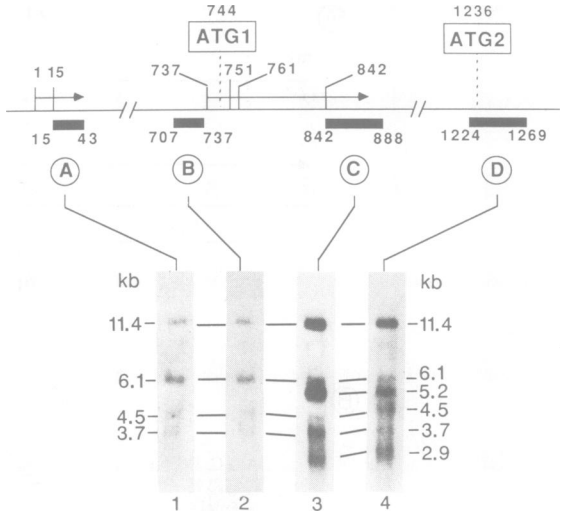


Fig. 6. Northern blot analysis of hPR mRNAs. Schematic illustration of the 5′ region of the hPR gene, showing the positions of ATG1, ATG2 and of the two clusters of transcriptional start sites (arrows), mapped in Figures 4 and 5. Below, synthetic oligonucleotides (A–D), complementary to the hPR mRNA, are depicted as bars; the position of their extremities is given. The bottom part of the figure shows the Northern blot, obtained by using ^{32}P-end-labelled oligonucleotides A–D as probes. In each lane, 7.5 μg of T47D poly(A)$^+$ RNA were loaded. On the left and right, the size of individual hPR transcripts estimated from DNA markers run in parallel are indicated. Note that a difference between the low molecular weight signals in lanes 4 and 5 is due to different times of electrophoretic separation (re-hybridization of 'blot C' with probe D showed the bands at the same position as in lane 3).

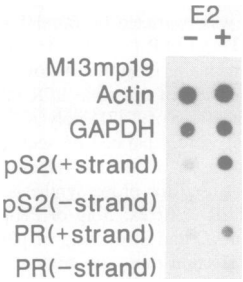


Fig. 7. Nuclear run-on transcription with MCF-7 nuclei demonstrates transcriptional induction by estrogens of the endogenous PR gene. Nuclei were isolated from estrogen-induced (+) or non-induced (–) MCF-7 cells, nascent trancripts were elongated *in vitro* in presence of ^{32}P-labelled CTP, and hybridized to filter-bound plasmid or M13 probes, as indicated. The PR probe originates from the M3–E4 gene fragment (Figure 1B); both strands were used to confirm strand-specificity of the signals (+ strand corresponds to faithful PR gene transcription). A pS2 gene probe was used as positive control of hormonal treatment; signals for actin and GADPH show that equal amounts of *in vitro* synthesized RNA were used for hybridization; M13mp19 gives the background signals for pS2 and PR probes due to vector hybridization.

site 2.3–2.4 kb further downstream may give rise to the 5.2 and 6.1 kb species. Note that the intensity of the 11.4 kb band (relative to, for example, the 6.1 kb band) was different for both sets of probes, suggesting that the 11.4 kb band detected with probes C and D may also correspond to two RNA species, initiating upstream and downstream of ATG1. In conclusion, the pattern of mRNA species observed with probes located upstream of +737 and downstream of +842 is fully compatible with the above S1 nuclease mapping data and suggests the existence of two promoter regions, B and A, generating transcripts encoding hPR forms B and A, respectively. Note that the 5.2 and 2.9 kb hPR RNAs (generating form A) are dominant species in T47D.

Induction of transcription of the hPR gene by estradiol in vivo

It has been shown previously that estrogen treatment of MCF-7 breast cancer cells increased hPR mRNA and protein

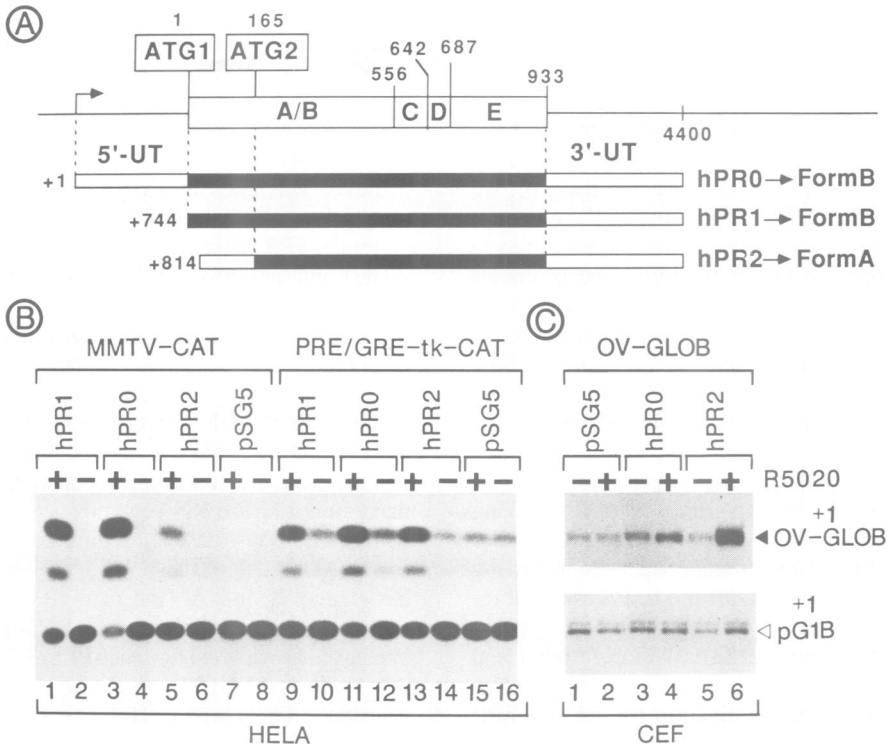


Fig. 8. Promoter-specific activation of transcription by the hPR forms A and B. (**A**) Schematic illustration of the expression vectors hPR0 and hPR1, generating form B, and hPR2, generating form A, as described (Kastner *et al.*, 1990). Solid boxes indicate the coding sequences, open boxes the 5′ and 3′ non-coding regions. At the top the modular structure of the hPR is depicted. (**B**) Progestin-dependent activation of transcription from the MMTV CAT (Cato *et al.*, 1986) (lanes 1–8) and PRE/GRE-*tk-CAT* (Green *et al.*, 1988b) (lanes 9–16) reporter recombinants by hPR forms A (expressed from hPR2) and B (expressed from hPR0 and hPR1). Reporter genes were transfected into HeLa cells at 1 μg, together with 100 ng (giving maximal stimulation, data not shown) of hPR expression vectors or pSG5, and 1 μg pCH110. Standardization of CAT assays was as described in Materials and methods. (**C**) Transcription from the OV-GLOB reporter gene in chicken embryo fibroblast (CEF) cells (Tora *et al.*, 1988a) is stimulated by hPR form A, but not form B, in the presence of progestins: 10 μg of OV-GLOB (+1 OV-GLOB signal) were transfected together with 1 μg of pSG5 (lanes 1 and 2), hPRO (lanes 3 and 4) or hPR2 (lanes 5 and 6) and 400 ng of the internal control pG1B (+1 pG1B signal). R5020 was added, where indicated, to a final concentration of 10^{-8} M. Cytoplasmic RNA was extracted and subjected to quantitative S1 nuclease mapping as described (Tora *et al.*, 1988a).

levels 3- to 10-fold and 40-fold, respectively (Nardulli *et al.*, 1988; Read *et al.*, 1988; Wei *et al.*, 1988). Moreover, a general increase of all PR mRNA species was noted (Wei *et al.*, 1988). To analyze whether this increase was a consequence of transcriptional stimulation by estradiol, *in vitro* run-on transcription assays were performed with nuclei isolated from stimulated (10 nM estradiol for 24 h) and non-treated MCF-7 cells ('+E2' and '−E2', respectively, in Figure 7). Hybridization of the [^{32}P]labelled *in vitro* extended transcripts to actin- and GAPDH-specific probes indicated that a similar amount of labelled RNA extracted from both sets of nuclei was used. The strand-specific signals obtained with a probe of the pS2 gene which is known to be induced by estrogens at the transcriptional level (Brown *et al.*, 1984) demonstrated that the hormonal treatment of MCF-7 cells was effective (note that 'M13mp19' gives the background signals obtained with single-stranded probes, Figure 7). Estrogen-inducibility of hPR gene transcription was analyzed with single-stranded probes originating from the 5′ part of the first intron [PR(+strand) and PR(−strand) in Figure 7]. Clearly, the amount of strand-specific *in vitro* extended nascent hPR gene transcripts was increased (5- to 6-fold) upon estrogen treatment (Figure 7). Exposure of cells to cycloheximide before hormonal treatment did not affect estrogen induction of PR gene transcription (data not shown). We thus concluded that hPR gene transcription in MCF-7 cells is induced by estrogens to a similar extent as reported for PR mRNA levels, suggesting the absence of significant estrogen-dependent effects on the stability of hPR mRNAs.

The hPR promoter regions A and B are estrogen inducible, but do not contain consensus palindromic estrogen responsive elements (EREs)

To investigate whether estrogen inducibility of hPR gene transcription is a characteristic of one or of both promoters, we transiently co-transfected the hER expression vector HEO (Green *et al.*, 1986) and chimeric genes bearing either of the two hPR promoter sequences into HeLa cells. The basal promoter activity of PR-[−711,+31]-CAT (Figure 4A) was stimulated 25-fold when transfected cells were exposed to estradiol (Figure 4C, compare lanes 2 and 3). Similar results were obtained when PR-[−711,+31]-GLOB (Figure 4A) was used in a quantitative S1 nuclease mapping analyses (Figure 4D, compare lanes 2 and 3). The three major transcription initiation sites (triangles) described above are clearly visible (the open triangles indicate transcription initiation at +1 and +15). Replacement of HEO with the parental expression vector pSG5 (Green *et al.*, 1988a) lacking the hER insert did not result in induction of transcription from the two chimeric genes, either in the absence or presence of estradiol (Figure 4C and D, lanes 4 and 5). Estrogen stimulation was also observed for the hPR promoter

region A, since PR-[464,1105]-CAT (Figure 5A) was estrogen/hER inducible in transient co-transfection assays (Figure 5C, compare lanes 1 and 2).

A survey of the sequences containing promoter A (+464 to +737) did not reveal the presence of a palindromic ERE (Klein-Hitpass *et al.*, 1987; Klock *et al.*, 1987; Martinez *et al.*, 1987; Berry *et al.*, 1989 and references therein); we noted, however, the 'half-palindromic' element +571-TGACC (boxed in Figure 2) which resembles a functional ERE mapped in the ovalbumin gene promoter (Tora *et al.*, 1988b). No consensus nor 'half-palindromic' ERE was detected in the region containing promoter B (−711 to +31). Further detailed promoter mapping will be required to define the regulatory elements that constitute the two estrogen responsive promoters of the hPR gene.

Forms A and B of hPR are functionally different

The apparent co-existence of two forms of hPR differing in their N-terminal region prompted us to investigate whether they might exhibit different functional properties. Knowing that the A/B region and the HBD of the hER contain transcriptional activation functions which exhibit a complex pattern of homo- and heterosynergism with each other and with upstream element factors (Tora *et al.*, 1989), we were curious to assess transcription activation by the two hPR forms of promoters bearing only one PRE/GRE and of more complex promoters with several progestin responsive elements, e.g. the mouse mammary tumor virus (MMTV) long terminal repeat (LTR) (Eul *et al.*, 1989 and references therein). Moreover, we previously observed different promoter specificities for the two forms of cPR, in that form A, but not form B, was able to activate the ovalbumin promoter, while both forms induced transcription from a MMTV LTR-based reporter gene (Tora *et al.*, 1988a).

Using expression vectors which selectively express either form of hPR (Kastner *et al.*, 1990; see Figure 8A) we obtained similar results to those mentioned above for the cPR; while in HeLa cells transcription from a reporter gene containing only one palindromic PRE/GRE (PRE/GRE-*tk*-CAT; Green *et al.*, 1988b) was activated similarly by both forms of hPR (Figure 8B, lanes 9, 11, 13), in chicken embryo fibroblasts (CEF), cell-specific transcription from the chicken ovalbumin promoter was stimulated only by hPR form A, as documented by quantitative S1 nuclease analysis (Figure 8C, compare lanes 3/4 and 5/6 relative to the signal from the internal control recombinant pG1B). Note in this respect the low degree of conservation of the N-terminal regions of cPR and hPR forms B (Figure 3). Furthermore, transcriptional activation by the hPR form B of the complex promoter in MMTV *CAT* (containing several 'imperfect' PRE/GREs), was 7-times stronger than that generated by form A (Figure 8B, compare lanes 1, 3 with lane 5). We conclude from these data that hPR forms A and B exert striking different promoter specificities, which suggests that the two forms have a different potential to synergize with one another and/or other upstream element factors involved in modulation of transcription.

Discussion

Two promoters generate distinct classes of transcripts encoding hPR forms A and B

The presence and origin of two forms of cPR and hPR (contrary to only one form described for rbPR, see Loosfelt *et al.*, 1984; Logeat *et al.*, 1985), has been a long-standing matter of controversy (see Conneely *et al.*, 1988; Gronemeyer, 1988; Horwitz and Francis, 1988 for reviews and references). It was generally agreed that cPR forms A and B result from initiation of translation at ATG128 and ATG1, respectively, but the mechanism by which this was achieved was unclear. Recently, it has been proposed that alternative internal initiation of translation at AUG128 generates cPR form A from a single mRNA containing also AUG1 and therefore that both forms A and B are translated from the same transcript (Conneely *et al.*, 1987b, 1989). We show here that in T47D breast cancer cells two promoters give rise to two distinct classes of hPR mRNAs. One of them directs initiation of transcription at +1 and +15, while the other promoter produces major transcripts with 5′ ends at +751, +761 and +842 which lack AUG1, and therefore are unable to encode hPR form B. It is unlikely that a minor hPR mRNA starting at +737 can code for form B, as the 5′ non-translated region is too short for translation from AUG1 (Kozak, 1987, 1989b). Both promoters, designated B and A (for form B and form A encoding transcripts), generate abundant mRNA species in T47D cells. We have shown recently that hPR mRNAs initiated between ATG1 and ATG165, give rise to only form A when transiently expressed in HeLa or Cos cells, while vectors generating transcripts containing AUG1 (and the 5′-untranslated region) produce exclusively form B (Kastner *et al.*, 1990). Collectively, these data argue against the hypothesis that form A is produced by alternative initiation of translation from mRNA transcribed from a single promoter. Rather, they strongly suggest that form A is generated from transcripts distinct from those encoding form B and that each of the promoters is responsible for the presence of only one of the hPR forms. Could a similar mechanism also generate the two forms of cPR? Most interestingly, we have recently detected in chick oviduct a cPR transcript that has the potential to encode only form A (Jeltsch *et al.*, 1990). However, while the presence of a cPR promoter, similar to hPR promoter B, could be demonstrated, the origin of the transcript encoding cPR form A remained unclear (Turcotte *et al.*, 1990).

No common feature could be detected by sequence comparison of the two hPR promoter regions A and B. Also, no sequence identities were found between hPR promoter B and its chicken counterpart; note, however, that none of the promoter regions contained a classical 'TATA' box (see Jeltsch *et al.*, 1990 for the cPR). On the contrary, a comparison between the 5′ gene flanking sequences of rabbit (Misrahi *et al.*, 1988) and human PRs revealed significant conservation (72% identity up to −711; sequences further upstream were not determined, see Figure 9). Remarkably, a putative Spl motif (5′-TGGGCGGGGC-3′, dashed in Figure 9; Kadonaga *et al.*, 1986) is embedded in a 46 bp long GC-rich peak (box in Figure 9, 75% G+C) which is well conserved in the two species at an identical distance from the two upstream cap sites, which are also conserved (Figure 9, arrows). In view of this sequence conservation, it is likely that rbPR gene transcription is also estrogen inducible, since estrogen is known to increase receptor mRNA concentration in rabbit uterus (Loosfelt *et al.*, 1984), while the absence of significant sequence identities between hPR and cPR (Jeltsch *et al.*, 1990) is in keeping with the lack of transcriptional regulation of the cPR gene by estrogen (see below and Turcotte *et al.*, 1990). With respect to

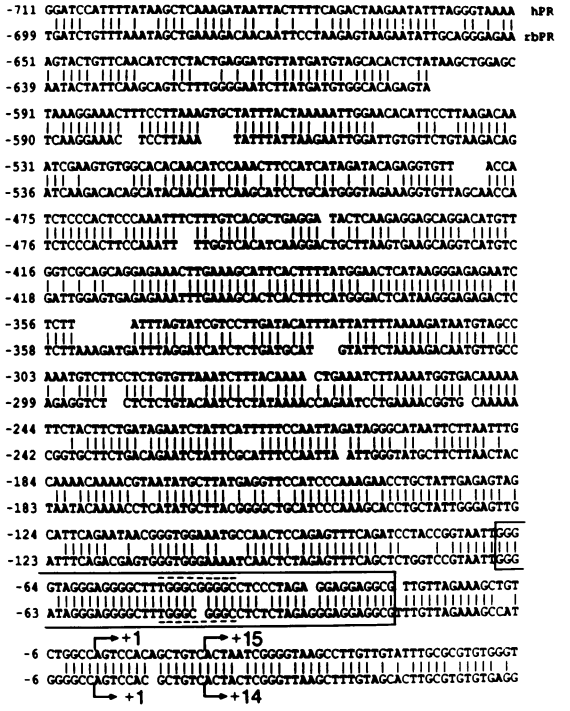


Fig. 9. Alignment of the 5′ gene flanking region of hPR (top sequence) and rbPR (bottom sequence). Positions relative to the most 5′ cap sites (+1) of hPR (this study) and rbPR (Misrahi *et al.*, 1988) are given at the left. A (G+C)-rich sequence close to the two cap sites (arrows), containing putative Sp1 binding sites (dashed lines) is boxed. Note that the positions of cap sites in hPR and rbPR has been conserved with respect to the alignment.

Fig. 10. Alignment of the 5′-untranslated region of hPR (top sequence) and rbPR (bottom sequence). Positions relative to the most 5′ cap site (+1) are depicted on the right. ATG codons are over- (hPR) or underlined (rbPR), the long ORFs initiate at position 744 (hPR) and 713 (rbPR, Misrahi *et al.*, 1988). Short UORFs (α and δ) present in the 5′ non-translated regions of the PR mRNAs are illustrated by dots, the corresponding ATG codons are over- or underlined. Dashes indicate potential Sp1 binding sites. A consecutive stretch of (T+C)- and (A+G)-rich motifs in front of ATG1, capable of forming a stem−loop structure ($\Delta G = -40$ kcal/mol) is schematically depicted by arrows, with the loop being indicated by an open box.

promoter A of hPR, we note the lack of one putative Sp1 binding site (dashed in Figure 10) in the corresponding sequence of the rbPR.

The human PR form B and rbPR mRNAs have highly conserved 5′-untranslated regions (74%, see Figure 10), whereas no significant sequence identity was found with the chicken homologue. Two UORFs (β and δ; in hPR, β is part of α) are highly conserved. Note that hPR UORF α initiates at an equivalent position in rbPR, but is shorter and lacks an internal ATG, while UORF γ is missing in the rabbit. UORF δ has a perfect 'Kozak' sequence for translational initiation (Kozak, 1989b). No UORF is present in mRNAs initiating from hPR promoter A. It remains to be established whether this reflects a differential control of translation for form B and form A mRNAs. The implication of UORFS in translational control has been documented for the yeast transcription factor GCN4 (Miller and Hinnebusch, 1989 and references therein). Note the presence in hPR form B mRNA of a sequence able to fold in a stable hairpin 13 bases in front of AUG1 (Figure 10) which has a shorter stem in the rabbit homologue. The potentially negative effect of such secondary RNA structures on translation *in vitro* has been demonstrated (Chevrier *et al.*, 1988; Kozak, 1989a).

Both hPR promoters are regulated by estrogen

Expression of the hPR gene in MCF-7 breast cancer cells can be stimulated by estrogen (Nardulli *et al.*, 1988; Read *et al.*, 1988). We have demonstrated here by using nuclear run-on transcription assays that the previously reported estrogen-dependent increase of hPR mRNA levels is mainly due to transcriptional stimulation. In keeping with these results we observed in transient transfection assays that both promoter A and promoter B are inducible by estrogen. While a 'half-palindromic' ERE might be involved in the estrogen responsiveness of promoter A, no candidate for an ERE could be found in promoter B sequence.

Estrogen regulation of PR gene expression in chick oviduct is, interestingly, strikingly different from that in human cells. We observed that in chick oviduct, estrogen-induction of cPR mRNA is a post-transcriptional event (Turcotte *et al.*, 1990 and references therein). It is unknown whether this is a tissue-specific effect or reflects a fundamental difference of regulation of PR expression by estrogens in human and chicken.

Possible physiological significance of the two PR isoforms

As discussed by Schibler and Sierra (1987) and Kozak (1988), transcription of a single gene from multiple promoters provides additional flexibility in the control of gene expression. It has been demonstrated that such promoters can have different cell type and/or development specific activity. Apart from hPR, only a few other cases have been described in which alternative start sites for transcription generate mRNAs with different coding potential (see Kozak, 1988 for references). The yeast invertase gene, for example, is transcribed from two promoters which produce functionally different enzymes (Carlson and Botstein, 1982). Interestingly, both promoters of the invertase gene are differently regulated, one being constitutively active, while the other is repressed at high, and

activated at low, glucose levels. Two antagonistic isoforms of the bovine papilloma virus E2 ORF have been described; one is a transcriptional activator, while the other represses transcription. Both forms apparently originate by translation of two in-frame AUGs of distinct mRNAs, generated from two promoters (Lambert *et al.*, 1987). It is unknown whether the activity of the two promoters A and B of hPR could be differentially regulated in a cell-specific manner. It is, however, interesting to note that in the presence of similar levels of estrogen receptors, estrogen-inducible and non-inducible progestin receptors were found in different areas of the rat brain (Romano *et al.*, 1989).

As previously demonstrated for cPR (Tora *et al.*, 1988a; Gronemeyer *et al.*, 1987), the two forms of hPR exhibit a strikingly different promoter specificity (see Figure 8) with three progestin-inducible promoters (for a further case, see Turcotte *et al.*, 1990). Both chicken and human PR forms A, but not B, were able to activate the ovalbumin promoter, while the MMTV promoter was activated more efficiently by form B than form A. However, the two hPR forms similarly activated PRE/GRE-*tk*-CAT which contains a monomeric responsive element. What might be the origin of these different promoter specificities? Since the MMTV contains multiple 'imperfect' PREs, the differences observed with the PRE/GRE-*tk*-CAT suggest that both PR forms have a different potential to synergize with each other or with the 'upstream element factors' present in these promoters. Synergistic effects of steroid hormone receptors on multiple responsive elements and with other transcription factors have been described recently (Ponglikitmongkol *et al.*, 1990; Tora *et al.*, 1989 and references therein). The reason for the inability of PR form B to activate the ovalbumin promoter is unclear. It may be related to a steric hindrance problem or reflect a more specific effect of the N-terminal region of hPR and cPR form B on the ovalbumin promoter activity. That hPR and cPR forms behave similarly appears to exclude a particular role of the polyglutamic acid region found in cPR form B (no such region is evident in hPR form B, see Figure 2). Whatever the mechanism of differential promoter activation may be, it is tempting to speculate that differential expression of hPR forms A and B from their cognate promoters, together with their differential target gene promoter specificities, might generate a system of tissue-specific regulation of progestin action. Further studies focussed on the expression of both hPR forms in various tissues are required to substantiate this hypothesis.

Materials and methods

Cloning and sequencing

The construction of the λgt11 T47D library has been described (Petkovich *et al.*, 1987). Screening was done with the nick-translated probes, under the conditions described by Walter *et al.* (1985). The human genomic library, established in λEMBL3 with partially *Sau*3A digested leucocyte DNA, was a gift of Transgène SA (Strasbourg).

The inserts to be sequenced were subcloned in both orientations into pEMBL18+. Sequencing was done on both strands using the Sanger dideoxy chain termination method. The reactions were performed at 55°C to minimize template secondary structures. Long inserts were sequenced by using the M13 universal primer and subclones generated by a DNase I deletion strategy (Lin *et al.*, 1985). Sequences of (G+C) rich regions were confirmed by the Maxam and Gilbert sequencing method (Ausubel *et al.*, 1987).

RNA isolation and Northern blotting

Total RNA was extracted from T47D cells with guanidine thiocyanate as described by Chirgwin *et al.* (1974). Poly(A)$^+$ RNA was purified by oligo(dT)−cellulose (Collaborative Research) chromatography (Maniatis *et al.*, 1982). Cytoplasmic RNA extracted from transfected cells was prepared by the Nonidet P-40 method (Groudine *et al.*, 1981).

Poly(A)$^+$ RNA was separated on formaldehyde−agarose gels and blotted onto nitrocellulose membranes (Schleicher and Schüll, BA85). Hybridization with end-labelled oligonucleotides (sp. act. 10^7 c.p.m./pmol) was performed at 42°C in 50% formaldehyde, 5 × SSPE (Maniatis *et al.*, 1982), 0.2 mg/ml salmon sperm DNA, 0.5 × Denhardt's and ~3−5 × 10^6 c.p.m. of probe. Washing was in 2 × SSPE, 0.1% SDS, at 25°C for 1 h and then for an additional hour in 0.2 × SSPE, 1% SDS at 60°C.

S1 nuclease mapping and primer extension

Single stranded S1 nuclease mapping probes were synthesized as follows: 5 pmol of end-labelled primer (Figure 4B; 5′-TTCTGCTGGCTCCGTA-CTG-3′; Figure 5A: 5′-AACAGACTGTCCAAGACACT-3′) were hybridized to 15−20 μg of cognate single-stranded DNA, elongated with the Klenow fragment of *Escherichia coli* DNA polymerase for 1 h in the presence of 0.25 mM deoxynucleotides, cut with the appropriate restriction enzyme and subsequently purified by electroelution from strand separating acrylamide gels (Maniatis *et al.*, 1982). The primer used for β-globin-based reporters is complementary to bases +39 to +60 of the rabbit β-globin gene (Zenke *et al.*, 1986). Hybridization was performed overnight with 2 × 10^4 c.p.m. of probe in a volume of 15 μl at 42°C in 50% formamide, 0.4 M NaCl, 40 mM PIPES, pH 6.5, 1 mM EDTA (experiments shown in Figure 4B and D) or at 55°C in 80% formamide, 0.4 M NaCl, 40 mM PIPES, pH 6.5, 1 mM EDTA (experiments shown in Figure 5B). S1 nuclease digestion was in 200 μl of 30 mM NaAc pH 4.5, 3 mM $ZnSO_4$, 300 mM NaCl containing 100 U of S1 nuclease (Appligène). Quantitative S1 nuclease analysis shown in Figure 8C was as described (Tora *et al.*, 1988a).

For primer extension, the primer (5′-AAGAGTAGCCCGTCCAGGG-AGATAGGTATGGCCGAAA-3′) was hybridized with poly(A)$^+$ RNA for 3 h at 55°C in 30 μl of 80% formamide, 0.4 M NaCl, 40 mM PIPES, pH 6.5, 1 mM EDTA. the annealed primer was then precipitated with ethanol and washed twice with 75% ethanol. Elongation was performed at 42°C for 1 h in 25 μl containing 1 mM dithiothreitol, 0.1 M Tris, pH 8.3, 10 mM $MgCl_2$, 1 mM deoxynculeotides, 1 U/μl RNasin (Promega) and 1 U/μl of AMV-reverse transcriptase ((Molecular Genetic Resources). The reaction was stopped by phenol/chloroform extraction and products were precipitated with ethanol prior to gel electrophoresis.

hPR expression vectors and chimeric genes

The human progesterone (hPR0, 1, 2, see Figure 8A) and estrogen (HEO) receptor expression vectors have been described (Kastner *et al.*, 1990; Green *et al.*, 1986). PR-[−711,+31]-CAT was constructed by creating a *Xho*I site at position +31 by site-directed mutagenesis (using the oligonucleotide 5′-GCAAATACCTCGAGGCTTACC-3′) and the *Bam*HI(−711)−*Xho*I(+31) fragment was inserted into the corresponding sites of pBLCAT8+ (Klein-Hitpass *et al.*, 1986, note that the *Xho*I site in pBLCAT8+ is downstream of the unique *Bgl*II site). PR-[−711,+31]-GLOB originates from TG2 (Meyer *et al.*, 1989), which was cut with *Bgl*II and an oligonucleotide, containing a *Xho*I/*Bam*HI/*Xba*I/*Ssp*I/*Bgl*II polylinker (PL in Figure 4A), was inserted. The resulting construct was cut with *Xho*I and *Hin*dIII to remove the *tk* promoter and the *Hin*dIII−*Xho*I fragment of PR-[−711,+31]-CAT was inserted PR-[+464,+1105]-CAT was constructed by ligating a *Mst*II/*Bgl*II linker to the *Bgl*II(+464)−*Mst*II(+1105) hPR cDNA fragment, digesting with *Bgl*II and inserting the fragment into *Bgl*II-cut pBLCAT3 (Lückow and Schütz, 1987).

Cell transfections

HeLa cells were transfected in Dulbecco's modified Eagles' medium (DMEM) with 5% fetal calf serum, treated with dextran-coated charcoal to remove steroid hormones. Cells used for studies involving estrogen regulation were cultivated in the absence of phenol red. Transfection was done at 50% confluency in 9 cm Petri dishes with a total amount of 20 μg supercoiled plasmid DNA using the calcium phosphate precipitation technique (Banerji *et al.*, 1981).

CAT assays

All transfections were performed with a mixture of plasmids which contained in addition to reporter and receptor expression vector 1 μg of the β-galactosidase expression vector pCH110 (Pharmacia) and Bluescribe M13+ (Stratagène) complementing to a total amount of 20 μg DNA. Hormones (see figure legends) were added at this time and 6−20 h post-transfection when the medium was changed. At 40 h post-transfection cells were

harvested in 1 ml/Petri dish ice-cold phosphate saline buffer (PBS), resuspended in 200 μl of 250 mM Tris, ph 7.5, lysed by three cycles of freeze−thaw, and centrifuged at 10 000 *g* for 15 min. β-galactosidase activity was measured as described by Herbomel *et al.* (1984). CAT assays were performed according to Gorman *et al.* (1982) with 0.1 μCi of [^{14}C]-chloramphenicol (50 Ci/mmol, Amersham) and amounts of extracts with equal β-galactosidase activities.

Nuclear run-on assays

Assays for nuclear run-on transcription were performed as described by Meyer *et al.* (1989), using hPR templates corresponding to either strand of the genomic M3−E4 fragment (Figure 1B). Southern blot analyses indicated the absence of repetitive sequences within this fragment (data not shown).

In vitro hPR RNA synthesis

hPRO RNA used as a control in Figure 5B was synthesized directly from hPRO linearized with *Hin*dIII using T7 polymerase as described (Kumar *et al.*, 1986).

Acknowledgements

We thank S.Green for the T47D cDNA library and PRE/GRE-*tk-CAT*, Transgène SA for the human leucocyte genomic library, and we gratefully acknowledge the excellent assistance of Isabelle Issemann. We thank S.Green, V.Kumar, A.Zelent and H.de Verneuil for advice, and J.M.Jeltsch, J.M.Garnier, M.E.Meyer and M.T.Bocquel for helpful discussions. We are grateful to A.Dierich and I.Scheuer for providing CEF cells, to the cell culture service for HeLa and Cos cells, and we thank the secretarial staff, as well as C.Werlé, B.Boulay and A.Landmann for help in preparing this manuscript. B.T. was a post-doctoral fellow of the Medical Research Council of Canada and is currently at the Massachusetts Institute of Technology, Department of Biology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA. U.S. is with Bayer, Leverkusen, FRG. This work was supported by the CNRS (A.I.V.), the INSERM (grants CNAMTS), the Ministère de la Recherche et de la Technologie, the Fondation pour la Recherche Médicale and the Association pour la Recherche sur le Cancer.

References

Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.G., Smith,J.A. and Struhl,K. (1987) In *Current Protocols in Molecular Biology*. J.Wiley, Chichester.

Banerji,J., Rusconi,S. and Schaffner,W. (1981) *Cell*, **27**, 299−687.

Beato,M. (1989) *Cell*, **56**, 335−344.

Berry,M., Nunez,A.M. and Chambon,P. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 1218−1222.

Brown,A.M.C., Jeltsch,J.M., Roberts,M. and Chambon,P. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 6344−6348.

Carlson,M. and Botstein,D. (1982) *Cell*, **28**, 145−154.

Cato,A.C.B., Miksicek,R., Schütz,G., Arnemann,J. and Beato,M. (1986) *EMBO J.*, **5**, 2237−2240.

Chevrier,D., Vézinat,C., Bastille,J., Linard,C., Sonenberg,N. and Boileau,G. (1988) *J. Biol. Chem.*, **263**, 902−910.

Chirgwin,J.M., Przybyla,A.E., McDonald,R.J. and Rutter,W.J. (1974) *Biochemistry*, **18**, 5284−5299.

Conneely,O.M., Dobson,A.D.W., Tsai,M.J., Beattie,W.G., Toft,D.O., Huckaby,C.S., Zarucki,T., Schrader,W.T. and O'Malley,B.W. (1987a) *Mol. Endocrinol.*, **1**, 517−525.

Conneely,O.M., Maxwell,B.L., Toft,D.O., Schrader,W.T. and O'Malley,B.W. (1987b) *Biophys. Res. Commun.*, **149**, 493−501.

Conneely,O.M., Kettelberger,D.M., Tsai,M.J., Schrader,W.T. and O'Malley,B.W. (1989) *J. Biol. Chem.*, **264**, 14062−14064.

Danielsen,M., Hinck,L. and Ringold,G.M. (1989) *Cell*, **57**, 1131−1138.

Eul,J., Meyer,M.E., Tora,L., Bocquel,M.T., Quirin-Stricker,C., Chambon,P. and Gronemeyer,H. (1989) *EMBO J.*, **8**, 83−90.

Evans,R.M. (1988) *Science*, **240**, 889−895.

Freedman,L.P., Luisi,B.F., Korszun,Z.R., Basavappa,R., Sigler,P.B. and Yamamoto,K.R. (1988) *Nature*, **334**, 543−546.

Godowski,P.J., Picard,D. and Yamamoto,K.R. (1988) *Science*, **241**, 812−816.

Gorman,C.M., Moffat,L.F. and Horward,B.H. (1982) *Mol. Cell. Biol.*, **2**, 1044−1051.

Green,S. and Chambon,P. (1988) *Trends Genet.*, **4**, 309−314.

Green,S., Walter,P., Kumar,V., Krust,A., Bornert,J.M., Argos,P. and Chambon,P. (1986) *Nature*, **320**, 134−139.

Green,S., Issemann,I. and Sheer,E. (1988a) *Nucleic Acids Res.*, **16**, 369.

Green,S., Kumar,V., Theulaz,I., Wahli,W. and Chambon,P. (1988b) *EMBO J.*, **7**, 3037−3044.

Gronemeyer,H. (1988) In Gronemeyer,H. (ed.), *Affinity Labelling and Cloning of Steroid and Thyroid Hormone Receptors. Ellis Horwood Series in Biomedicine*. VHC Publishers, New York, pp. 167−185.

Gronemeyer,H., Turcotte,B., Quirin-Stricker,C., Bocquel,M.T., Meyer,M.E., Krozowski,Z., Jeltsch,J.M., Lerouge,T., Garnier,J.M. and Chambon,P. (1987) *EMBO J.*, **6**, 3985−3994.

Gronemeyer,H., Turcotte,B., Quirin-Stricker,C., Bocquel,M.T., Meyer, M.E., Krozowski,Z., Jeltsch,J.M., Lerouge,T., Garnier,J.M. and Chambon,P. (1988) In Gronemeyer,H. (ed.), *Affinity Labelling and Cloning of Steroid and Thyroid Hormone Receptors. Ellis Horwood Series in Biomedicine*. VCH Publishers, New York, pp. 252−297.

Groudine,M., Peretz,M. and Weintraub,H. (1981) *Mol. Cell. Biol.*, **1**, 281−288.

Ham,J. and Parker,M.G. (1989) *Curr. Opinion Cell Biol.*, **1**, 503−511.

Herbomel,P., Bourachot,B. and Yaniv,M. (1984) *Cell*, **39**, 653−662.

Hollenberg,S.M .and Evans,R.M. (1988) *Cell*, **55**, 899−906.

Horwitz,K.B. and Francis,M.D. (1988) In Gronemeyer,H. (ed.), *Affinity Labelling and Cloning of Steroid and Thyroid Hormone Receptors. Ellis Horwood Series in Biomedicine*. VCH Publishers, New York, pp. 186−198.

Jeltsch,J.M., Krozowski,Z., Quirin-Stricker,C., Gronemeyer,H., Simpson,R.J., Garnier,J.M., Krust,A., Jacob,F. and Chambon,P. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 5424−5428.

Jeltsch,J.M., Turcotte,B., Garnier,J.M., Lerouge,T., Krozowski,Z., Gronemeyer,H. and Chambon,P. (1990) *J. Biol. Chem.*, in press.

Kadonaga,J.T., Jones,K.A. and Tjian,R. (1986) *Trends Biochem. Sci.*, **2**, 20−23.

Kastner,P., Bocquel,M.T., Turcotte,B., Garnier,J.M., Horwitz,K.B., Chambon,P. and Gronemeyer,H. (1990) submitted.

Klein-Hitpass,L., Ryffel,G.U., Heilinger,E. and Cato,A. (1987) *Nucleic Acids Res.*, **16**, 647−663.

Klock,G., Strähle,U. and Schütz,G. (1987) *Nature*, **329**, 734−736.

Kozak,M. (1987) *Nucleic Acids Res.*, **15**, 8125−8148.

Kozak,M. (1988) *J. Cell. Biol.*, **107**, 1−7.

Kozak,M. (1989a) *Mol. Cell. Biol.*, **9**, 5134−5142.

Kozak,M. (1989b) *J. Cell. Biol.*, **108**, 229−241.

Krett,N.L., Wei,L.L., Francis,M.D., Nordeen,S.K., Gordon,D.F., Wood,W.M. and Horwitz,K.B. (1988) *Biochem. Biophys. Res. Commun.*, **157**, 278−285.

Krust,A., Green,S., Argos,P., Kumar,V., Walter,P., Bornert,J.M. and Chambon,P. (1986) *EMBO J.*, **5**, 891−897.

Kumar,V., Green,S., Staub,A. and Chambon,P. (1986) *EMBO J.*, **5**, 2231−2236.

Kumar,V. and Chambon,P. (1988) *Cell*, **55**, 145−156.

Lambert,P.F., Spalholz,B.A. and Howley,P.M. (1987) *Cell*, **50**, 69−78.

Lin,H.C., Lei,S.P. and Wilcox,G. (1985) *Anal. Biochem.*, **147**, 114−119.

Logeat,F., Pamphile,R., Loosfelt,H., Jolivet,H., Fournier,A. and Milgrom,E. (1985) *Biochemistry*, **24**, 1029−1035.

Loosfelt,H., Logeat,F., Hai,M.T.V. and Milgrom,E. (1984) *J. Biol. Chem.*, **259**, 14196−14202.

Loosfelt,H., Atger,M., Misrahi,M., Guiochon-Mantel,A., Meriel,C., Logeat,F., Benarous,R. and Milgrom,E. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 9045−9049.

Lückow,B. and Schütz,G. (1987) *Nucleic Acids Res.*, **15**, 5490.

Mader,S., Kumar,V., de Verneuil,H. and Chambon,P. (1989) *Nature*, **338**, 271−274.

Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Martinez,E., Givel,F. and Wahli,W. (1987) *EMBO J.*, **6**, 3719−3727.

Meyer,M.E., Gronemeyer,H., Turcotte,B., Bocquel,M.T., Tasset,D. and Chambon,P. (1989) *Cell*, **57**, 433−442.

Miller,P.F. and Hinnebusch,A.G. (1989) *Genes Dev.*, **3**, 1217−1225.

Misrahi,M., Atger,M., d'Auriol,L., Loosfelt,H., Meriel,C., Fridlanski,F., Guiochon-Mantel,A., Galibert,F. and Milgrom,E. (1987) *Biochem. Biophys. Res. Commun.*, **143**, 740−748.

Misrahi,M., Loosfelt,H., Atger,M., Mériel,C., Zerah,V., Dessen,P. and Milgrom,E. (1988) *Nucleic Acids Res.*, **16**, 5459−5472.

Nardulli,A.M., Greene,G.L., O'Malley,B.W. and Katzenellenbogen,B.S. (1988) *Endocrinology*, **122**, 935−943.

Petkovich,M., Brand,N.J., Krust,A. and Chambon,P. (1987) *Nature*, **330**, 444−450.

Ponglikitmongkol,M., White,J.H. and Chambon,P. (1990) submitted.

Read,L.D., Snider,C.E., Miller,J.S., Greene,G.L. and Katzenellenbogen,S. (1988) *Mol. Endocrinol.*, **2**, 263−271.

Romano,G.J., Krust,A. and Pfaff,D.W. (1989) *Mol. Endocrinol.*, **3**, 1295–1300.
Sassone-Corsi,P., Duboule,D. and Chambon,P. (1985) *Cold Spring Harbor Symp. Quant. Biol.*, **50**, 747–752.
Schibler,U. and Sierra,F. (1987) *Annu. Rev. Genet.*, **21**, 237–257.
Tora,L., Gronemeyer,H., Turcotte,B., Gaub,M.P. and Chambon,P. (1988a) *Nature*, **333**, 185–188.
Tora,L., Gaub,M.P., Mader,S., Dierich,A., Bellard,M. and Chambon,P. (1988b) *EMBO J.*, **7**, 3771–3778.
Tora,L., White,J., Brou,C., Tasset,D., Webster,N., Scheer,E. and Chambon,P. (1989) *Cell*, **59**, 477–487.
Turcotte,B., Meyer,M.E., Bellard,M., Dretzen,G., Gronemeyer,H. and Chambon,P. (1990) submitted.
Umesono,K. and Evans,R.M. (1989) *Cell*, **57**, 1139–1146.
Walter,P., Green,S., Greene,G., Krust,A., Bornert,J.M., Jeltsch,J.M., Staub,A., Jensen,E., Scrace,G., Waterfield,M. and Chambon,P. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 7889–7893.
Webster,N.J.G., Green,S., Jin,J.R. and Chambon,P. (1988) *Cell*, **54**, 199–207.
Webster,N.J.G., Green,S., Tasset,D., Ponglikitmongkol,M. and Chambon,P. (1989) *EMBO J.*, **8**, 1441–1446.
Wei,L.L., Krett,N.L., Francis,M.D., Gordon,D.F., Wood,W.M., O'Malley,B.W. and Horwitz,K.B. (1988) *Mol. Endocrinol.*, **2**, 62–72.
Zenke,M., Grundström,T., Matthes,H., Wintzerith,M., Schatz,C., Wildeman,A. and Chambon,P. (1986) *EMBO J.*, **5**, 387–397.

Received on January 12, 1990